

by F.A. Ausubel et al., John Wiley and Sons, Inc. (1987). Stringency conditions can be determined readily by the skilled artisan. An example of moderate stringency hybridization conditions would be hybridization in 5X SSC, 5X Denhardt's Solution, 50% (w/v) formamide, and 1% SDS at 42°C with washing conditions of 0.2X SSC and 0.1% SDS at 42°C. An example of high stringency conditions can be defined as hybridization conditions as above, and with washing at approximately 68°C, in 0.1X SSC, and 0.1% SDS. The skilled artisan will recognize that the temperature and wash solution salt concentration can be adjusted as necessary according to factors such as the length of the probe.

[054] Due to the degeneracy of the genetic code wherein more than one codon can encode the same amino acid, multiple DNA sequences can code for the same polypeptide. Such variant DNA sequences can result from genetic drift or artificial manipulation (e.g., occurring during PCR amplification or as the product of deliberate mutagenesis of a native sequence). The present invention thus encompasses any nucleic acid capable of encoding a protein derived from SEQ ID NOS: 1-327, or variants thereof.

[055] Deliberate mutagenesis of a native sequence can be carried out using numerous techniques well known in the art. For example, oligonucleotide-directed site-specific mutagenesis procedures can be employed, particularly where it is desired to mutate a gene such that predetermined restriction nucleotides or codons are altered by substitution, deletion or insertion. Exemplary methods of making such alterations are disclosed by Walder et al. (*Gene* 42:133, 1986); Bauer et al. (*Gene* 37:73, 1985); Craik (*BioTechniques*, January 12-19, 1985); Smith et al. (*Genetic Engineering: Principles*

and Methods, Plenum Press, 1981); Kunkel (*Proc. Natl. Acad. Sci. USA* 82:488, 1985); Kunkel et al. (*Methods in Enzymol.* 154:367, 1987); and U.S. Patent Nos. 4,518,584 and 4,737,462, all of which are incorporated by reference.

[056] Thus, the invention further provides an isolated nucleic acid molecule selected from the group comprising of (1), (2), and (3) above and further consisting of:

- (4) an isolated nucleic acid molecule degenerate from SEQ ID NOS: 1-334 as a result of the genetic code; and
- (5) an isolated nucleic acid molecule selected from the group consisting of allelic variants and species homologs of SEQ ID NOS: 1-334.

Obtaining Full Length cDNAs

[057] The cDNAs isolated and cloned through the differential display procedure will often only represent a partial sequence (generally the 3' end) of the mRNA from which it was derived due to the nature of the arbitrary primer used in the differential display PCR reaction. Consequently, the cDNA sequences of SEQ ID NOS: 1-327 provide powerful tools for obtaining the sequences of full-length cDNAs. This can be accomplished by using a partial cDNA as a probe to identify and isolate the full length cDNA from a population of full length cDNAs or from a full length cDNA library. As is well known in the art, similar procedures can be used to identify corresponding genomic DNA sequences.

[058] Alternatively, one can obtain the 5' sequence of a partial cDNA by performing Rapid Amplification of cDNA Ends (RACE) procedures such as those disclosed in Frohman, *Methods in Enzymology*, 218:340-356 (1993) and Bertling et al., *PCR Methods and Applications* 3:95-99 (1993) which are hereby incorporated by

INNEGAN, HENDERSON,
FARABOW, GARRETT,
& DUNNER, L.L.P.
1300 I STREET, N.W.
WASHINGTON, DC 20005
202-408-4000

reference. For example, Clonetech Laboratories, Inc. (Palo Alto, CA) offers a SMART™ cDNA product line that allows one to generate high quality full length cDNAs and cDNA libraries. SMART™ technology can also be used to perform RACE. One skilled in the art will readily recognize that there are other equivalent products and procedures for obtaining full length cDNAs. Full length cDNAs may be sequenced and their sequences compared to sequences in public databases to assess their identities and/or homologies to other known sequences.

[059] Cloned full length cDNAs can be used in the construction of expression vectors for the production and purification of pine tree polypeptides which contain the pine tree peptides encoded by the cDNAs of any one of SEQ ID NOS: 1-327.

[060] Oligonucleotide Primers for PCR Assays

[061] In another embodiment, the present invention encompasses oligonucleotide fragments derived from any one of SEQ ID NO: 1 through SEQ ID NO: 327 or from the reverse complement sequence of any one of SEQ ID NO: 1 through SEQ ID NO: 327. Such oligonucleotides would be useful as primers in the performance of RT-PCR assays to detect, or even quantify, pine embryo stage-specific transcripts. Such oligonucleotide primers will generally comprise from 10 to 25 nucleotides substantially complementary to the ends of the target sequence and may contain additional non-complementary nucleotides, for example, nucleotides that generate a restriction endonuclease site or cloning junction. Programs useful in selecting PCR primers may be used to design the oligonucleotides of this invention, but use of such programs is not necessary. By way of example, the Wisconsin Package™ software available from the Genetic Computer Group (Madison, Wisconsin) includes a program